

## **STIC Biotechnology Systems Branch**

### **RAW SEQUENCE LISTING** **ERROR REPORT**

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/591,632C  
Source: IFW16  
Date Processed by STIC: 2/7/06

**THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.**

**PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:**

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

**FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221**

**TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.4.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:**

**<http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm>**

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

- 1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)**
- 2. U.S. Postal Service:** Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
- 3. Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 01/14/05):**  
U.S. Patent and Trademark Office, Mail Stop Sequence, Customer Window, Randolph Building, 401 Dulany Street, Alexandria, VA 22314

Revised 01/10/06

## Raw Sequence Listing Error Summary

<u>ERROR DETECTED</u>	<u>SUGGESTED CORRECTION</u>	SERIAL NUMBER: <u>09/591,632C</u>
ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE		
1 <u>    </u> Wrapped Nucleics Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor <b>after</b> creating it. Please adjust your right margin to .3; this will prevent "wrapping."	
2 <u>    </u> Invalid Line Length	The rules require that a line <b>not exceed</b> 72 characters in length. This includes white spaces.	
3 <u>    </u> Misaligned Amino Numbering	The numbering under each 5 <sup>th</sup> amino acid is misaligned. Do <b>not</b> use tab codes between numbers; use <b>space characters</b> , instead.	
4 <u>    </u> Non-ASCII	The submitted file was <b>not</b> saved in ASCII(DOS) text, as <b>required</b> by the Sequence Rules. Please <b>ensure your subsequent submission is saved in ASCII text</b> .	
5 <u>    </u> Variable Length	Sequence(s) <u>        </u> contain n's or Xaa's representing more than one residue. <b>Per Sequence Rules, each n or Xaa can only represent a single residue.</b> Please present the <b>maximum</b> number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.	
6 <u>    </u> PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) <u>        </u> . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. <b>This applies to the mandatory &lt;220&gt;-&lt;223&gt; sections for Artificial or Unknown sequences.</b>	
7 <u>    </u> Skipped Sequences (OLD RULES)	Sequence(s) <u>        </u> missing. If intentional, please insert the following lines for <b>each</b> skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped  Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to <b>include</b> the skipped sequences.	
8 <u>    </u> Skipped Sequences (NEW RULES)	Sequence(s) <u>        </u> missing. If <b>intentional</b> , please insert the following lines for <b>each</b> skipped sequence. <210> sequence id number <400> sequence id number 000	
9 <u>    </u> Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is <b>MANDATORY</b> if n's or Xaa's are present. In <220> to <223> section, please explain location of <b>n</b> or <b>Xaa</b> , and which residue <b>n</b> or <b>Xaa</b> represents.	
10 <u>    </u> Invalid <213> Response	Per 1.823 of Sequence Rules, the only <b>valid</b> <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is <b>required</b> when <213> response is Unknown or is Artificial Sequence	
11 <u>    </u> Use of <220>	Sequence(s) <u>        </u> missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is <b>MANDATORY</b> if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)	
12 <u>    </u> PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.	
13 <u>    </u> Misuse of n/Xaa	<b>"n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid</b>	



IFW16

## RAW SEQUENCE LISTING

DATE: 02/07/2006

PATENT APPLICATION: US/09/591,632C

TIME: 09:01:23

Input Set : A:\34978a.txt

Output Set: N:\CRF4\02072006\I591632C.raw

3 <110> APPLICANT: Lindquist, et al.  
 5 <120> TITLE OF INVENTION: RECOMBINANT PRION-LIKE GENES AND PROTEINS AND MATERIALS AND  
 6 METHODS COMPRISING SAME  
 8 <130> FILE REFERENCE: 30554/34978A  
 C--> 10 <140> CURRENT APPLICATION NUMBER: US/09/591,632C  
 C--> 10 <141> CURRENT FILING DATE: 2000-06-09  
 10 <150> PRIOR APPLICATION NUMBER: US 09/591,632  
 11 <151> PRIOR FILING DATE: 2000-06-09  
 13 <150> PRIOR APPLICATION NUMBER: US 60/138,833  
 14 <151> PRIOR FILING DATE: 1999-06-09  
 16 <160> NUMBER OF SEQ ID NOS: 70  
 18 <170> SOFTWARE: PatentIn version 3.3

*Use of records  
 72 characters*

*(see item 2 on*

*Error*

*summary*

*sheet)*

*delete -  
 there are  
 not prior  
 data. They  
 are current data.*

## ERRORED SEQUENCES

429 <210> SEQ ID NO: 3  
 430 <211> LENGTH: 1427  
 431 <212> TYPE: DNA  
 432 <213> ORGANISM: Saccharomyces cerevisiae  
 434 <220> FEATURE:  
 435 <221> NAME/KEY: CDS  
 436 <222> LOCATION: (182)..(1246)  
 438 <400> SEQUENCE: 3  
 439 ctctgaggttg aaaagaatag caaaaatctt tccttttcaa acagctcatt tggaattggt 60  
 441 tatagcactg aattgaatcg aagaggaata aagatcccc gtacgaactt ctttatTTTT 120  
 443 agtttttcat tttttgttat tagtcatatt gttttaagct gcaaattaag ttgtacacca 180  
 445 a atg atg aat aac aac ggc aac caa gtg tcg aat ctc tcc aat gcg ctc 229  
 446 Met Met Asn Asn Asn Gly Asn Gln Val Ser Asn Leu Ser Asn Ala Leu  
 447 1 5 10 15  
 449 cgt caa gta aac ata gga aac agg aac agt aat aca acc acc gat caa 277  
 450 Arg Gln Val Asn Ile Gly Asn Arg Asn Ser Asn Thr Thr Thr Asp Gln  
 451 20 25 30  
 453 agt aat ata aat ttt gaa ttt tca aca ggt gta aat aat aat aat aat 325  
 454 Ser Asn Ile Asn Phe Glu Phe Ser Thr Gly Val Asn Asn Asn Asn Asn  
 455 35 40 45  
 457 aac aat agc agt agt aat aac aat aat gtt caa aac aat aac agc ggc 373  
 458 Asn Asn Ser Ser Ser Asn Asn Asn Asn Val Gln Asn Asn Asn Ser Gly  
 459 50 55 60  
 461 cgc aat ggt agc caa aat aat gat aac gag aat aat atc aag aat acc 421  
 462 Arg Asn Gly Ser Gln Asn Asp Asn Glu Asn Asn Ile Lys Asn Thr  
 463 65 70 75 80  
 465 tta gaa caa cat cga caa caa caa cag gca ttt tcg gat atg agt cac 469

**Does Not Comply  
 Corrected Diskette Needed**

*pp 1, 3, 5-6, 7*

## RAW SEQUENCE LISTING

DATE: 02/07/2006

PATENT APPLICATION: US/09/591,632C

TIME: 09:01:23

Input Set : A:\34978a.txt

Output Set: N:\CRF4\02072006\I591632C.raw

466	Leu	Glu	Gln	His	Arg	Gln	Gln	Gln	Gln	Ala	Phe	Ser	Asp	Met	Ser	His	
467					85					90					95		
469	gtg	gag	tat	tcc	aga	att	aca	aaa	ttt	ttt	caa	gaa	caa	cca	ctg	gag	517
470	Val	Glu	Tyr	Ser	Arg	Ile	Thr	Lys	Phe	Phe	Gln	Glu	Gln	Pro	Leu	Glu	
471				100					105					110			
473	gga	tat	acc	ctt	ttc	tct	cac	agg	tct	gcg	cct	aat	gga	ttc	aaa	gtt	565
474	Gly	Tyr	Thr	Leu	Phe	Ser	His	Arg	Ser	Ala	Pro	Asn	Gly	Phe	Lys	Val	
475			115					120					125				
477	gct	ata	gta	cta	agt	gaa	ctt	gga	ttt	cat	tat	aac	aca	atc	ttc	cta	613
478	Ala	Ile	Val	Leu	Ser	Glu	Leu	Gly	Phe	His	Tyr	Asn	Thr	Ile	Phe	Leu	
479		130					135					140					
481	gat	ttc	aat	ctt	ggc	gaa	cat	agg	gcc	ccc	gaa	ttt	gtg	tct	gtg	aac	661
482	Asp	Phe	Asn	Leu	Gly	Glu	His	Arg	Ala	Pro	Glu	Phe	Val	Ser	Val	Asn	
483	145				150					155					160		
485	cct	aat	gca	aga	gtt	cca	gct	tta	atc	gat	cat	ggg	atg	gac	aac	ttg	709
486	Pro	Asn	Ala	Arg	Val	Pro	Ala	Leu	Ile	Asp	His	Gly	Met	Asp	Asn	Leu	
487				165					170					175			
489	tct	att	tgg	gaa	tca	ggg	gcg	att	tta	tta	cat	ttg	gta	aat	aaa	tat	757
490	Ser	Ile	Trp	Glu	Ser	Gly	Ala	Ile	Leu	Leu	His	Leu	Val	Asn	Lys	Tyr	
491			180					185					190				
493	tac	aaa	gag	act	ggg	aat	cca	tta	ctc	tgg	tcc	gat	gat	tta	gct	gac	805
494	Tyr	Lys	Glu	Thr	Gly	Asn	Pro	Leu	Leu	Trp	Ser	Asp	Asp	Leu	Ala	Asp	
495			195				200					205					
497	caa	tca	caa	atc	aac	gca	tgg	ttg	ttc	ttc	caa	acg	tca	ggg	cat	gcg	853
498	Gln	Ser	Gln	Ile	Asn	Ala	Trp	Leu	Phe	Phe	Gln	Thr	Ser	Gly	His	Ala	
499		210				215					220						
501	cca	atg	att	gga	caa	gct	tta	cat	ttc	aga	tac	ttc	cat	tca	caa	aag	901
502	Pro	Met	Ile	Gly	Gln	Ala	Leu	His	Phe	Arg	Tyr	Phe	His	Ser	Gln	Lys	
503	225				230					235					240		
505	ata	gca	agt	gct	gta	gaa	aga	tat	acg	gat	gag	gtt	aga	aga	gtt	tac	949
506	Ile	Ala	Ser	Ala	Val	Glu	Arg	Tyr	Thr	Asp	Glu	Val	Arg	Arg	Val	Tyr	
507			245					250					255				
509	ggg	gta	gtg	gag	atg	gcc	ttg	gct	gaa	cgt	aga	gaa	gcg	ctg	gtg	atg	997
510	Gly	Val	Val	Glu	Met	Ala	Leu	Ala	Glu	Arg	Arg	Glu	Ala	Leu	Val	Met	
511			260					265					270				
513	gaa	tta	gac	acg	gaa	aat	gcg	gct	gca	tac	tca	gct	ggg	aca	aca	cca	1045
514	Glu	Leu	Asp	Thr	Glu	Asn	Ala	Ala	Ala	Tyr	Ser	Ala	Gly	Thr	Thr	Pro	
515			275				280					285					
517	atg	tca	caa	agt	cgt	ttc	ttt	gat	tat	ccc	gta	tgg	ctt	gta	gga	gat	1093
518	Met	Ser	Gln	Ser	Arg	Phe	Phe	Asp	Tyr	Pro	Val	Trp	Leu	Val	Gly	Asp	
519		290				295				300							
521	aaa	tta	act	ata	gca	gat	ttg	gcc	ttt	gtc	cca	tgg	aat	aat	gtc	gtg	1141
522	Lys	Leu	Thr	Ile	Ala	Asp	Leu	Ala	Phe	Val	Pro	Trp	Asn	Asn	Val	Val	
523	305				310					315					320		
525	gat	aga	att	ggc	att	aat	atc	aaa	att	gaa	ttt	cca	gaa	gtt	tac	aaa	1189
526	Asp	Arg	Ile	Gly	Ile	Asn	Ile	Lys	Ile	Glu	Phe	Pro	Glu	Val	Tyr	Lys	
527			325						330					335			
529	tgg	acg	aag	cat	atg	atg	aga	aga	ccc	gcg	gtc	atc	aag	gca	ttg	cgt	1237
530	Trp	Thr	Lys	His	Met	Met	Arg	Arg	Pro	Ala	Val	Ile	Lys	Ala	Leu	Arg	

## RAW SEQUENCE LISTING

DATE: 02/07/2006

PATENT APPLICATION: US/09/591,632C

TIME: 09:01:23

Input Set : A:\34978a.txt

Output Set: N:\CRF4\02072006\I591632C.raw

```

531          340          345          350
533 ggt gga tga aggtctgttt aaaaacaaga aagaaagaag aaggaggaaa      1286
534 Gly Gly
537 agaaggttat aagggtatgt atataggcag acaaaaagga aaattaagtg caaatataaa      1346
539 caaaaatgtc atagaagtat ataatagttt tgaaatttct gttgcttcta tttatttctt      1406
E--> 541 gttaccccaa ccacagaatt      1427/1426
724 <210> SEQ ID NO: 11
725 <211> LENGTH: 446 445
726 <212> TYPE: DNA
727 <213> ORGANISM: Artificial sequence
729 <220> FEATURE:
730 <223> OTHER INFORMATION: CUP1 promoter
732 <400> SEQUENCE: 11
733 ccattaccga catttggcgc ctatacgtgc atatgttcat gtatgtatct gtatttaaaa      60
735 cacttttcta ttatttttcc tcatatatgt gtataggttt atacggatga tttaattatt      120
737 acttcaccac cctttatttc aggtctgatc cttagccttg ttactagtta gaaaaagaca      180
739 tttttgctgt cagtcactgt caagagattc ttttgctggc atttcttcta gaagcaaaaa      240
741 gagcgatgcg tcttttccgc tgaaccgttc cagcaaaaaa gactaccaac gcaatatgga      300
743 ttgtcagaat catataaaaag aagaagcaaa taactccttg tcttgtatca attgcattat      360
745 atatcttctt gttagtgcga tatcatatag aagtcatcga aatagatatt aagaaaaaca      420
E--> 747 aactgtacaa tcaatcaatc aatca      445
3712 <210> SEQ ID NO: 45
3713 <211> LENGTH: 7239 7238 (p 5-6)
3714 <212> TYPE: DNA
3715 <213> ORGANISM: Artificial sequence
3717 <220> FEATURE:
3718 <223> OTHER INFORMATION: Vector containing chimeric gene
3720 <400> SEQUENCE: 45
3721 gacgaaaggg cctcgtgata cgcctatttt tataggttaa tgtcatgata ataatggttt      60
3723 cttaggacgg atcgcttgcc tgtaacttac acgcgcctcg tatcttttaa tgatggaata      120
3725 atttgggaat ttactctgtg tttatttatt tttatgtttt gtatttggat tttagaaagt      180
3727 aaataaaagaa ggtagaagag ttacggaatg aagaaaaaaa aataaacaac ggtttaaaaa      240
3729 atttcaacaa aaagcgtact ttacatatat atttattaga caagaaaagc agattaaata      300
3731 gatatacatt cgattaacga taagtaaaat gtaaaatcac aggattttcg tgtgtggtct      360
3733 tctacacaga caagatgaaa caattcggca ttaataacct agagcaggaa gagcaagata      420
3735 aaaggtagta tttgttggcg atccccctag agtcttttac atcttcggaa aacaaaaaact      480
3737 attttttctt taattttctt ttttactttc tatttttaat ttatatattt atattaaaaa      540
3739 atttaaatta taattatttt tatagcacgt gatgaaaagg acccaggtgg cacttttcgg      600
3741 ggaaatgtgc gcggaacccc tatttgttta tttttctaaa tacattcaaa tatgtatccg      660
3743 ctcatgagac aataaccctg ataaatgctt caataatatt gaaaaaggaa gagtatgagt      720
3745 attcaacatt tccgtgtcgc ccttattccc ttttttgcgg cattttgcct tccgtttttt      780
3747 gctcaccag aaacgctggt gaaagtaaaa gatgctgaag atcagttggg tgcacgagtg      840
3749 ggttacatcg aactggatct caacagcggg aagatccttg agagttttcg ccccgagaa      900
3751 cgttttccaa tgatgagcac ttttaaagt tctgctatgt gcgcggtatt atcccgatt      960
3753 gacgccgggc aagagcaact cggtcgccgc atacactatt ctcaaatga cttggttgag      1020
3755 tactcaccag tcacagaaaa gcatcttac gatggcatga cagtaagaga attatgcagt      1080
3757 gctgccataa ccatgagtga taacactcgc gccaaacttac ttctgacaac gatcggagga      1140
3759 ccgaaggagc taaccgcttt tttgcacaac atgggggatc atgtaactcg ccttgatcgt      1200
3761 tgggaaccgg agctgaatga agccatacca aacgacgagc gtgacaccac gatgcctgta      1260

```

## RAW SEQUENCE LISTING

DATE: 02/07/2006

PATENT APPLICATION: US/09/591,632C

TIME: 09:01:23

Input Set : A:\34978a.txt

Output Set: N:\CRF4\02072006\I591632C.raw

3763	gcaatggcaa	caacgttgcg	caaaactatta	actggcgaaac	tacttactct	agcttccccg	1320
3765	caacaattaa	tagactggat	ggaggcggat	aaagttgcag	gaccacttct	gcgctcggcc	1380
3767	cttccggctg	gctggtttat	tgctgataaa	tctggagccg	gtgagcgtgg	gtctcgcggt	1440
3769	atcattgcag	cactggggcc	agatggtaag	ccctcccgtg	tcgtagttag	ctacacgacg	1500
3771	gggagtcagg	caactatgga	tgaacgaaat	agacagatcg	ctgagatagg	tgccctactg	1560
3773	attaagcatt	ggtaactgtc	agaccaagtt	tactcatata	tacttttagat	tgatttaaaa	1620
3775	cttcattttt	aatttaaaag	gatctagggtg	aagatccttt	ttgataatct	catgaccaaa	1680
3777	atcccctaac	gtgagttttc	gttccactga	gcgtcagacc	ccgtagaaaa	gatcaaagga	1740
3779	tcttcttgag	atcctttttt	tctgcgcgta	atctgctgct	tgcaaacaaa	aaaaccaccg	1800
3781	ctaccagcgg	tggtttgttt	gccggatcaa	gagctaccaa	ctctttttcc	gaaggtaact	1860
3783	ggcttcagca	gagcgcagat	accaaatact	gtccttctag	tgtagccgta	gttaggccac	1920
3785	cacttcaaga	actctgtagc	accgcctaca	tacctcgtc	tgctaactct	gttaccagt	1980
3787	gctgctgcc	gtggcgataa	gtcgtgtctt	accgggttgg	actcaagacg	atagttaaccg	2040
3789	gataaggcgc	agcggtcggg	ctgaacgggg	ggttcgtgca	cacagcccag	cttgagcgca	2100
3791	acgacctaca	ccgaactgag	atacctacag	cgtgagctat	gagaaagcgc	cacgcttccc	2160
3793	gaagggagaa	aggcggacag	gtatccggta	agcggcaggg	tcggaacagg	agagcgcacg	2220
3795	aggagagctt	cagggggaaa	cgcttggtat	ctttatagtc	ctgtcgggtt	tcgccacctc	2280
3797	tgacttgagc	gtcgattttt	gtgatgctcg	tcaggggggc	ggagcctatg	gaaaaacgcc	2340
3799	agcaacgcgg	cctttttacg	gttccctggc	ttttgctggc	cttttgctca	catgttcttt	2400
3801	cctgcgttat	cccctgatcc	tgtggataac	cgtattaccg	cctttgagt	agctgatacc	2460
3803	gctcgccgca	gccgaacgac	cgagcgcagc	gagtcagtga	gcgaggaagc	ggaagagcgc	2520
3805	ccaatacgca	aaccgcctct	ccccgcgcgt	tgcccgattc	attaatgcag	ctggcacgac	2580
3807	aggtttcccg	actggaaagc	gggcagtgag	cgcaacgcaa	ttaatgtgag	ttacctcact	2640
3809	cattaggcac	cccaggcttt	acactttatg	cttccggctc	gtatgttgtg	tggaattgtg	2700
3811	agcggataac	aatttcacac	aggaaacagc	tatgaccatg	attacgcaa	gctcgggaatt	2760
3813	aacctcact	aaagggaaac	aaagctgggt	accgggcccc	ccctcgaggt	cgacgggtatc	2820
3815	gataagcttg	atatcgaatt	cccattaccg	acatttgggc	gctatacgtg	catatgttca	2880
3817	tgtatgtatc	tgtatttaaa	acacttttgt	attatttttc	ctcatatatg	tgtataggtt	2940
3819	tatacgtagt	atttaattat	tacttcacca	ccctttattt	caggctgata	tcttagcctt	3000
3821	gttactagtt	agaaaaagac	atttttgctg	tcagtcactg	tcaagagatt	cttttgctgg	3060
3823	catttcttct	agaagcaaaa	agagcgtatg	gtcttttccg	ctgaaccgtt	ccagcaaaaa	3120
3825	agactaccaa	cgcaatatgg	attgtcagaa	tcatataaaa	gagaagcaaa	taactccttg	3180
3827	tcttgtatca	attgcattat	aatatcttct	tgtagtgca	atatcatata	gaagtcacg	3240
3829	aaatagatat	taagaaaaac	aaactgtaca	atcaatcaat	caatcaggat	ccatggatac	3300
3831	ggataagtta	atctcagagg	ctgagtctca	tttttctcaa	ggaaaccatg	cagaagctgt	3360
3833	tgcaaggttg	acatccgcag	ctcagtcgaa	ccccaatgac	gagcaaatgt	caactattga	3420
3835	atcattaatt	caaaaaatcg	caggatacgt	catggacaac	cgtagtgggtg	gtagtgcgc	3480
3837	ctcgcaagat	cgtgctgctg	gtgggtgggtc	atcttttatg	aacactttta	tgccagactc	3540
3839	taagggttct	tcccaaacgc	aactaggaaa	actagctttg	ttagccacag	tgatgacaca	3600
3841	ctcatcaaa	taaagggttct	ctaacagagg	gtttgacgta	gggactgtca	tgtcaatgct	3660
3843	aagtgggtct	ggcggcggga	gccaaagtat	gggtgcttcc	ggcctggctg	ccttggcttc	3720
3845	tcaattcttt	aagtcaggta	acaattccca	aggtcaggga	caaggtcaag	gtcaagggtca	3780
3847	aggtcaagga	caaggtcaag	gtcaagggtc	ttttactgct	ttggcgtctt	tggttctatc	3840
3849	tttcatgaat	tccaacaaca	ataatcagca	aggtcaaaat	caaagctccg	gtggttcttc	3900
3851	ctttggagca	ctagcttcta	tggcaagttc	tttatgcat	tccaataata	atcagaactc	3960
3853	caacaatagt	caacaggggt	ataaccaatc	ctatcaaaac	ggtaaccaaa	atagtcaagg	4020
3855	ttacaataat	caacagtacc	aagggtggcaa	cgggtggttac	caacaacaac	agggacaatc	4080
3857	tggtgggtgct	ttttctcat	tggcctccat	ggctcaatct	tacttaggtg	gtggacaaac	4140
3859	tcaatccaac	caacagcaat	acaatcaaca	aggccaaaac	aaccagcagc	aataccagca	4200

## RAW SEQUENCE LISTING

DATE: 02/07/2006

PATENT APPLICATION: US/09/591,632C

TIME: 09:01:23

Input Set : A:\34978a.txt

Output Set: N:\CRF4\02072006\I591632C.raw

group  
9  
↑

3861	acaaggccaa	aactatcagc	accaacaaca	gggtcagcag	cagcaacaag	gccactccag	4260
3863	ttcatttctca	gctttggctt	ccatggcaag	ttcctacctg	ggcaataact	ccaattcaaa	4320
3865	ttcgagttat	gggggccagc	aacaggctaa	tgagtatggt	agaccacaac	acaatggtca	4380
3867	acaacaatct	aatgagtacg	gaagaccgca	atacggcgga	aaccagaact	ccaatggaca	4440
E--> 3869	gcacgaatcc	cttaattttt	ctggcaactt	ttctcaacag	aacaataacg	gcaaccagaa	4500 4499
E--> 3871	ccgctaccgg	cggatggcta	gcaaaggaga	agaactcttc	actggagttg	tcccaattct	4560
E--> 3873	tggtgaatta	gatggtgatg	ttaatgggca	caaattttct	gtcagtggag	aggggtgaagg	4620
E--> 3875	tgatgcaaca	tacggaaaac	ttacccttaa	atatttttgc	actactggaa	aactacctgt	4680
E--> 3877	tccatggcca	acacttgctca	ctactttcac	ttatggtggt	cagtgtcttt	caagataccc	4740
E--> 3879	ggatcatatg	aaacggcatg	actttttcaa	gagtgccatg	cccgaagggt	atgtacagga	4800
E--> 3881	aagaactata	tttttcaaag	atgacgggaa	ctacaagaca	cgtgctgaag	tcaagtttga	4860
E--> 3883	aggtgatacc	cttgtaata	gaatcgagtt	aaaagggtatt	gattttaaag	aagatggaaa	4920
E--> 3885	cattcttggg	cacaaattgg	aatacaacta	taactcacac	aatgtataca	tcatggcaga	4980
E--> 3887	caaacaaaag	aatggaatca	aagctaactt	caaaattaga	cacaacattg	aagatggaaag	5040
E--> 3889	cgttcaacta	gcagaccatt	atcaacaaaa	tactccaatt	ggcgatggcc	ctgtcctttt	5100
E--> 3891	accagacaac	cattacctgt	ccacacaatc	tgccctttcg	aaagatccca	acgaaaagag	5160
E--> 3893	agaccacatg	gtccttcttg	agtttgtaac	agctgctggg	attacacatg	gcatggatga	5220
E--> 3895	actatacaaa	tgagagctcc	aattcgccct	atagtgaatc	gtattacaat	tcactggccg	5280
E--> 3897	tcgtttttaca	acgtcgtgac	tgggaaaacc	ctggcggtac	ccaacttaat	cgccttgacg	5340
E--> 3899	cacatccccc	tttcgccagc	tgggcgaata	gcgaagaggc	ccgcaccgat	cgccttccc	5400
E--> 3901	aacagttgag	cagcctgaat	ggcgaatggc	gcgacgcgcc	ctgtagcggc	gcattaagcg	5460
E--> 3903	cggcggtgtg	ggtggttacg	cgcagcgtga	ccgctacact	tgccagcgcc	ctagcgccc	5520
E--> 3905	ctccttttcgc	tttcttccct	tcctttctcg	ccacgttcgc	cggctttccc	cgtaagctc	5580
E--> 3907	taaatcgggg	gctcccttta	gggttccgat	ttagtgtctt	acggcacctc	gaccccaaaa	5640
E--> 3909	aacttgatta	gggtgatggt	tcacgtatgt	ggccatcgcc	ctgatagacg	gtttttcgcc	5700
E--> 3911	ctttgacgtt	ggagtccacg	ttctttaata	gtggactctt	gttccaaact	ggaacaacac	5760
E--> 3913	tcaaccctat	ctcgggtctat	tcttttgatt	tataagggat	tttgccgatt	tcggcctatt	5820
E--> 3915	ggttaaaaaa	tgagctgatt	taacaaaaat	ttaacgcgaa	ttttaacaaa	atattaacgt	5880
E--> 3917	ttacaatttc	ctgatgcggt	attttctcct	tacgcaatct	tgccggtatt	cacaccgat	5940
E--> 3919	agggtataata	ctgatataat	taaattgaag	ctctaatttg	tgagttagt	atacatgcat	6000
E--> 3921	ttacttataa	tacagttttt	tagttttgct	ggccgcatct	tctcaaatat	gcttcccagc	6060
E--> 3923	ctgctttttct	gtaacgttca	ccctctacct	tagcatccct	tccctttgca	aatagtcctc	6120
E--> 3925	ttccaacaat	aataatgtca	gatcctgtag	agaccacatc	atccacgggt	ctatactgtt	6180
E--> 3927	gacccaatgc	gtctcccttg	tcacttaaac	ccacaccggg	tgctcataatc	aaccaatcgt	6240
E--> 3929	aaccttcac	tcttccaccc	atgtctcttt	gagcaataaa	gccgataaca	aaatctttgt	6300
E--> 3931	cgtctcttcgc	aatgtcaaca	gtacccttag	tatattctcc	agtagatagg	gagcccttgc	6360
E--> 3933	atgacaattc	tgctaacatc	aaaaggcctc	taggttcctt	tgttacttct	tctgccgcct	6420
E--> 3935	gcttcaaacc	gctaacaata	cctgggcca	ccacaccgtg	tgcatctgta	atgtctgcc	6480
E--> 3937	attctgctat	tctgtataca	cccgcagagt	actgcaattt	gactgtatta	ccaatgtcag	6540
E--> 3939	caaattttct	gtcttcgaag	agtaaaaaat	tgtacttggc	ggataatgcc	tttagcggct	6600
E--> 3941	taactgtgcc	ctccatggaa	aaatcagtca	agatatccac	atgtgttttt	agtaaacaaa	6660
E--> 3943	ttttgggacc	taatgcttca	actaactcca	gtaattcctt	ggtggtacga	acatccaatg	6720
E--> 3945	aagcacacaa	gtttgtttgc	ttttcgtgca	tgatattaaa	tagcttggca	gcaacaggac	6780
E--> 3947	taggatgagt	agcagcacgt	tccttatatg	tagctttcga	catgatttat	cttcgtttcc	6840
E--> 3949	tgagggtttt	tgttctgtgc	agttgggtta	agaatactgg	gcaatttcac	gtttcttcaa	6900
E--> 3951	cactacatat	gcgtatatat	accaaatctaa	gtctgtgctc	cttccttcgt	tcttcttctc	6960
E--> 3953	gttcggagat	taccgaatca	aaaaaatttc	aaagaaaccg	aaatcaaaaa	aaagaataaa	7020
E--> 3955	aaaaaaatga	tgaattgaat	tgaagagctg	tggtatgggtg	cactctcagt	acaatctgct	7080
E--> 3957	ctgatgccgc	atagttaagc	cagccccgac	acccgccaac	acccgctgac	gcgccctgac	7140

number  
ff

## RAW SEQUENCE LISTING

DATE: 02/07/2006

PATENT APPLICATION: US/09/591,632C

TIME: 09:01:23

Input Set : A:\34978a.txt

Output Set: N:\CRF4\02072006\I591632C.raw

E--> 3959 gggcttgtct gctccccggca tccgcttaca gacaagctgt gaccgtctcc gggagctgca  
E--> 3961 tgtgtcagag gttttcaccg tcatcaccga aacgcgcga

7200  
7239

hos.  
off



from sequence 39

09/591,632C

7

gta aca gct gct ggg att aca cat ggc atg gat gaa cta tac aaa tga 720  
Val Thr Ala Ala Gly Ile Thr His Gly Met Asp glu Leu Tyr Lys  
225 230 235

(240) delete, since  
no amino  
acid  
is shown

VERIFICATION SUMMARY

DATE: 02/07/2006

PATENT APPLICATION: US/09/591,632C

TIME: 09:01:24

Input Set : A:\34978a.txt

Output Set: N:\CRF4\02072006\I591632C.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application No  
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:541 M:254 E: No. of Bases conflict, LENGTH:Input:1427 Counted:1426 SEQ:3 ✓  
L:541 M:252 E: No. of Seq. differs, <211> LENGTH:Input:1427 Found:1426 SEQ:3 ✓  
L:747 M:252 E: No. of Seq. differs, <211> LENGTH:Input:446 Found:445 SEQ:11 ✓  
L:3593 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:39 ✓  
L:3869 M:254 E: No. of Bases conflict, LENGTH:Input:4500 Counted:4499 SEQ:45 ✓  
M:254 Repeated in SeqNo=45  
L:3961 M:252 E: No. of Seq. differs, <211> LENGTH:Input:7239 Found:7238 SEQ:45 ✓